



Early Journal Content on JSTOR, Free to Anyone in the World

This article is one of nearly 500,000 scholarly works digitized and made freely available to everyone in the world by JSTOR.

Known as the Early Journal Content, this set of works include research articles, news, letters, and other writings published in more than 200 of the oldest leading academic journals. The works date from the mid-seventeenth to the early twentieth centuries.

We encourage people to read and share the Early Journal Content openly and to tell others that this resource exists. People may post this content online or redistribute in any way for non-commercial purposes.

Read more about Early Journal Content at <http://about.jstor.org/participate-jstor/individuals/early-journal-content>.

JSTOR is a digital library of academic journals, books, and primary source objects. JSTOR helps people discover, use, and build upon a wide range of content through a powerful research and teaching platform, and preserves this content for future generations. JSTOR is part of ITHAKA, a not-for-profit organization that also includes Ithaka S+R and Portico. For more information about JSTOR, please contact support@jstor.org.

VARIABILITY AS RELATED TO SEX DIFFERENCES IN ACHIEVEMENT

A CRITIQUE

LETA STETTER HOLLINGWORTH
New York City

This paper is the outcome of prolonged reflection on the doctrine of greater male variability. It comprises an attempt to assemble and review briefly data at present accessible as to the comparative variability of the sexes in mental traits, and to discuss critically the hypothesis that the great difference between the sexes in intellectual achievement and eminence is due to the inherently greater variability of the males. This hypothesis is stated clearly and concisely by Thorndike¹ thus:

The trivial difference between the central tendency of men and that of women which is the common finding of psychological tests and school experience may seem at variance with the patent fact that in the great achievements of the world in science, art, invention, and management, women have been far excelled by men. One who accepts the equality of typical (i.e., modal) representatives of the two sexes, must assume the burden of explaining this great difference in the high ranges of achievement.

The probably true explanation is to be sought in the greater variability within the male sex. . . .

In particular, if men differ in intelligence and energy by wider extremes than do women, eminence in and leadership of the world's affairs of whatever sort will inevitably belong oftener to men. They will oftener deserve it.

It is at once evident how important are the implications here stated for those who hope much from the present tendency to remove all disabilities of law, custom, and prejudice from women. If the explanation of women's failure to achieve significant things in the fields named by Thorndike is really to be found in the inherently greater variability of males, then complete liberation of women from excessive maternity and from all the consequent customs and legal disabilities that have developed, will result

¹ E. L. Thorndike, *Educational Psychology* (1910), p. 35.

only in raising the *average* intelligence and happiness of the race. We shall not expect any increase from this source in the number of eminent individuals, nor in achievement of that high order which forces knowledge and wisdom farther.

Thorndike¹ states the implications for pedagogy thus:

This one fundamental difference in variability is more important than all the differences between the average male and female capacities . . . a slight excess of male variability would mean that of the hundred most gifted individuals in this country not two would be women, and of the thousand most gifted, not one in twenty. . . . Women may and doubtless will be scientists and engineers, but the Joseph Henry, the Rowland, and the Edison of the future will be men; even should all women vote, they would play a small part in the Senate. . . . Not only the probability and the desirability of marriage and the training of children as an essential feature of woman's career, *but also the restriction of women to the mediocre grades of ability and achievement should be reckoned with by our educational systems.* The education of women for such professions as administration, statesmanship, philosophy, or scientific research, where a very few gifted individuals are what society requires, is far less needed than education for such professions as nursing, teaching, medicine, or architecture, where the average level is the essential. . . . Postgraduate instruction, to which women are flocking in large numbers is, at least in its higher reaches, a far more remunerative investment in the case of men.²

The first discussion of the comparative variability of the sexes bore on anatomical traits, and began about a century ago. The anatomist Meckel³ concluded on pathological grounds that the human female showed greater variability than the human male, "and he thought that since man is the superior animal and variation a sign of inferiority, the conclusion was justified." Later, when anatomists and naturalists arrived at the conclusion that the male is more variable, variability came to be regarded as an advantage, a characteristic affording the greatest hope for progress, and finally as the probable explanation of the fact that all the world's greatest deeds of intellect have been the deeds of men. This latter view obtains at present among men of science, though not without exceptions, the most notable of whom is Karl Pearson.⁴

¹ E. L. Thorndike, "Sex in Education," *The Bookman*, XXIII, 213.

² The italics here are mine.

³ Meckel, *Manual of Descriptive and Pathological Anatomy* (see Ellis, *Man and Woman* [1909], p. 410).

⁴ Karl Pearson, *Chances of Death* (1897).

It will be well at this point to consider not only the social and biological significance of variability, but also the connotation of the term itself, and whether every author who discusses variability means the same thing. There is, in fact, complaint among authors that the term is indefinite. Even in their controversial matter,¹ Ellis and Pearson complain of each other that there is failure to define the word. Theoretically greater variability always implies *greater range*, if the trait distributed conforms to the Gauss curve of probability. Empirical data, however, are not yet forthcoming to demonstrate that mental traits conform to the theoretical curve; and there is at present no conclusive empirical evidence to show that in cases where the coefficient of variation is greater for one sex than for the other, this greater variability consists in *greater range*. If we neglect theory and confine ourselves to facts as demonstrated, greater variability is found to consist in any or all of three typical conditions:

1. Greater range (Series B as compared with Series A).
2. Equal range for both groups, but greater frequency at the extremes for one group (Series C as compared with Series A).
3. Smaller range for the more variable group, with slight flattening at the top of the curve of distribution (Series D as compared with Series A).

A fourth condition is found in the work of Bonser, where the males are seen to be more variable than the females, though the *range* for the sexes is equal, and the *frequency at both extremes* is nearly twice as great for the females. This case will be taken up later in connection with other results from Bonser.

Let us now consider a hypothetical case. Table I gives four possible distributions of the same trait, including the same number of cases. This trait may be, for example, ability to perform an amount of work in a specified time, this ability being indicated by units varying from 1 to 15. Let Series A be a group of 1,000 women, and let Series B, C, and D be groups of 1,000 men each. It is seen that these Series all show greater variability on the part of the males (reference to Table I will show just how much greater is the *A.D.* in each case), but the social implications differ widely.

¹ H. Ellis, *Man and Woman* (Appendix).

In Series B the greater variability of the males consists in greater *range*. It is on this Series that we might base the explanation of the fact that *all* the world's greatest deeds of intellect have been the deeds of men; for here no women equal the best men.

In Series C the greater variability of the males consists in greater frequency at the extremes, *the range being equal*. On this Series might be based an explanation of the fact that *more* men than women have reached the *same* degree of eminence. It would not explain why no women have reached the greatest eminence.

In Series D the greater variability of the males consists in a flattening at the top of the curve of distribution, the *range* for the men being actually *less* than for the women.

Now it is clear that if social significance is to be attached to greater variability, not only the coefficient of variation must be stated, but also *what form the distribution takes*. Obviously a greater male variability like that shown in Series D would have no validity at all in explaining why the greatest deeds of intellect have been the deeds of men. If greater male variability takes this form, all the greatest deeds will be those of *women*.

In his discussions of greater male variability and its implications for pedagogy, Thorndike¹ theoretically means greater range: "Though the central tendencies were the same, there would still be two men of the hundred who were better than the best woman and two men who were worse than the worst woman." This condition would be represented under Series B. But, in discussing certain statistics regarding third-year high-school classes see Table I, p. 514.

This condition would be that of Series C. The range for the sexes is equal, but the frequency at the extremes is greater for males. Such cases of greater variability do not suggest an explanation of the fact that no women have achieved the greatest intellectual eminence. They would only explain the condition in which twice as many men as women achieved the *same* intellectual eminence. But our chief problem is to explain why no women have equaled the best men.

Havelock Ellis,² in a chapter on "The Variational Tendency

¹ E. L. Thorndike, *op. cit.*, p. 42.

² H. Ellis, *op. cit.*, p. 412.

of Men," discusses certain anatomical and pathological data which show, on the whole, the greater variability of the male. Karl Pearson, in a polemical article, undertook to disprove the conclusions of Ellis, stigmatizing them as "scientific superstition." This controversy between Ellis and Pearson is very familiar to students of social science, and each of us may weigh the evidence for himself, since we have here two authorities, of perhaps equal competence, in diametrical disagreement.

TABLE I

| DEGREE OR AMOUNT OF THE TRAIT MEASURED | FREQUENCY | FREQUENCY | FREQUENCY | FREQUENCY |
|---|-----------|-----------|-----------|-----------|
| | Series A | Series B | Series C | Series D |
| 1..... | 0 | 1 | 0 | 0 |
| 2..... | 0 | 3 | 0 | 0 |
| 3..... | 1 | 7 | 6 | 0 |
| 4..... | 10 | 14 | 18 | 11 |
| 5..... | 45 | 42 | 60 | 57 |
| 6..... | 117 | 115 | 112 | 130 |
| 7..... | 205 | 200 | 197 | 190 |
| 8..... | 244 | 236 | 214 | 224 |
| 9..... | 205 | 200 | 197 | 190 |
| 10..... | 117 | 115 | 112 | 130 |
| 11..... | 45 | 42 | 60 | 57 |
| 12..... | 10 | 14 | 18 | 11 |
| 13..... | 1 | 7 | 6 | 0 |
| 14..... | 0 | 3 | 0 | 0 |
| 15..... | 0 | 1 | 0 | 0 |

| | SERIES A Women | SERIES B Men | SERIES C Men | SERIES D Men |
|------------------------|-------------------|-----------------|-----------------|-----------------|
| (Standard group) | | | | |
| Number..... | = 1,000 | = 1,000 | = 1,000 | = 1,000 |
| Central Tendency..... | = 8 | = 8 | = 8 | = 8 |
| Average Deviation..... | = 1.238 | = 1.544 | = 1.406 | = 1.330 |

On the whole boys are twice as frequent as girls in the youngest and oldest age, and about one and one-half times as frequent at ages fourteen and nineteen.

But if it were definitely proved that there is greater male variability in anatomical measurements, it would only suggest, not prove, that there is greater male variability in *mental* traits also. Very, very little precise evidence has been adduced as to the comparative variability of the sexes in mental traits. Such general evidence as that previously brought forward, for instance

by Ellis, that the great geniuses of the world have been men, and that there are at the same time more idiots among men, is obviously fallacious. For the geniuses on the one hand may be accounted for by the fact that woman's biological function of reproduction has so conditioned her that eminence in the fields where mental energy is publicly recognized has been extremely improbable; and we should expect statisticians to find more idiots and feeble-minded individuals among men, because they take their data from institutions, where defective men are more likely to be admitted than women of the same degree of defectiveness. Women have been and are a dependent and non-competitive class, and when defective can more easily survive outside of institutions, since they do not have to compete *mentally* with normal individuals, as men do, to maintain themselves in the social *milieu*. This conclusion is well confirmed by the records of the Clearing-House for Mental Defectives at the Post-Graduate Hospital in New York City. Among 1,000 consecutive cases of mental defect (including idiocy, imbecility, and feeble-mindedness), taken from all cases diagnosed at this Clearing-House during the years 1912 and 1913, there were 568 males and 432 females. But of individuals *over sixteen years of age* there were only 78 males, while there were 159 females; and of individuals *over 30 years of age* there were 9 males and 28 females. A detailed account of this study may be found in an article recently published.¹ At present it suffices to point out that the fact that females escape the Clearing-House till beyond the age of thirty years three times as frequently as males, fits very well with the fact that more males than females are brought to the Clearing-House, on the whole. The boy who cannot compete mentally is found out, becomes at an early age an object of concern to relatives, is brought to the Clearing-House, and directed toward an institution. The girl who cannot compete mentally is not so often recognized as definitely defective, since it is not unnatural for her to drop into the isolation of the home, where she can "take care of" small children, peel potatoes, scrub, etc. If physically passable, as is often the case, she may marry,

¹ L. S. Hollingworth, "The Frequency of Amentia as Related to Sex," *Medical Record*, 1913.

thus fastening herself to economic support; or she may become a prostitute, to which economic pursuit feeble mentality is no barrier. Thus they survive *outside of institutions*. The writer has frequently questioned those who accompany these feeble-minded women over thirty years of age to the Clearing-House. Their tardy appearance there is usually accounted for by the fact that some accident has at last happened: "her husband has just died"; "she has rheumatism, and can scrub no more"; "an illegitimate pregnancy has again befallen, to the distraction of relatives"; "she was a prostitute, but physical illness has driven her in from the street." No one can doubt that there are scores of feeble-minded women at large, to whom these accidents have not happened.

It will be well at this point to survey and compare precise data already at hand to show sex differences in mental variability. Such data have been assembled here from scattered sources. Thorndike¹ gives precise data tending to show greater mental variability in men and boys. He calculated as well as might be from data given, the variability for each sex in the traits tested by Helen Bradford Thompson.² His results show that men are about one-twentieth more variable than women, in these experiments. He also concludes from certain measurements of reaction time, spelling, arithmetical ability, etc., that "it is extremely probable that, except in the two years nearest the age of puberty for girls, the male sex is slightly more variable."

Wissler's results with college students show female variability to be in general about nine-tenths that of males. The number of women measured was, however, only 42, and the ratio of female variability differed greatly in the different traits, so that the nine-tenths would, by itself alone, be of no great reliability.³

Thorndike deplores the fact that there is so little precise data at hand, but leaves us to suppose that he considers what is available as sufficient to lend a very, very high degree of probability to the conclusion which he states, and which was quoted at the outset. Several articles and monographs, however, have appeared

¹ E. L. Thorndike, *op. cit.*, pp. 33-43.

² H. B. Thompson, *The Mental Traits of Sex* (1906).

³ Thompson, *op. cit.*

since 1910 which are in disagreement with the results cited by this author.

Wells¹ in a study of "Sex Differences in the Tapping Test" reached the conclusion that men are more variable than women. He had, however, only ten subjects, five women and five men—too small a number on which to generalize. In another study including five women and five men this author concludes:

The groups of subjects are perhaps too small to expect any special sex differences to be illustrated. . . . In the addition test the performance of the women is much more variable than that of the men, in the number-checking test it is much less so.

H. L. Hollingworth² made a study of judgments of persuasiveness, using advertisements as material. He had as subjects 20 Juniors in Barnard College and 20 Juniors in Columbia College. Among his conclusions he states: "Men correlate with their group average about 25 per cent more closely than women," and "the range of variability in the above coefficient is for the men only 43 per cent as large as for women." In the course of discussion this author says:

Another set of measurements of interest is found in the figures which show the approximation of the individual's judgments to the average judgment of his group. . . . The coefficients for the women range between -0.13 and 0.66 , thus giving a total range of 0.79 , with the average at 0.48 . For the men the coefficients cover a much narrower range, varying between 0.40 and 0.74 , thus giving a total range of only 0.34 , a range only 43 per cent as large as that of the women. The average for the men is 0.59 , the median is 0.61 , being thus about 25 per cent higher than the same for the women. Only four women exceed the median for men, while all the men but four exceed the median for women.

Both of these facts—that of higher correlation and that of narrower range—point in the same direction, that is, toward the greater homogeneity of the group of men. The high coefficients indicate that any one man selected at random will be a better example of the characteristics of his group than will a similarly selected woman of her group. And the narrow range again indicates the tendency of the men, not only to depart but slightly from the type, but also to depart in approximately equal degrees from it. Whether

¹ F. L. Wells, "Sex Differences in the Tapping Test," *American Journal of Psychology*, 1909.

² H. L. Hollingworth, "Judgments of Persuasiveness," *Psychological Review* XXVIII (1911), 4.

these facts point to a greater general variability of women as compared with men, or only to the particular composition of the two groups taking part in this experiment, one cannot say. But the present method seems to indicate a concrete and interesting way of studying this much disputed question of the relative variability of the sexes, in what may be called the higher mental processes.

E. K. Strong, Jr.,¹ in a study of the merits of advertisements by the method of relative position, had twenty-five subjects—fifteen men and ten women. Among his conclusions he states the following:

An inspection of the diagram of Table I shows that the range of judgments for the men is much less than for the women, i.e., from $+0.84$ to 0 for the men, and from $+0.75$ to -0.43 for the women. Both have 55 per cent of the entire range below the median judgment. But the average *A.D.* of the medians of the individual judgments for each advertisement for the women is 69 per cent greater than for the men. This is the more striking, as the women would apparently be a more homogeneous group than the men, as they were all Juniors or Seniors in Barnard College, and within a very few years of each other in age, while the men included graduate students and professors and vary at least twenty years in age. . . . A comparison of the two groups shows us that the *P.E.* of the women averages 69.7 per cent greater than that of the men.

In the arrangements of another series of advertisements, where a greater number of subjects was used, this author found the women to be less variable than the men. He remarks upon these contradictory findings as follows:

It is true that the methods employed in the two chapters are different. But if different methods can give exactly opposite results as to variability, they can be of little value as to its determination. Personally I believe that the situation is this. The results of chap. vii show that when women are given an equal opportunity with men to rate appeals (advertisements) they are able to classify their dislikes as readily as their preferences, which the men do not do. Such a condition naturally results in a greater total range where methods of experimentation similar to those in this chapter are used, and consequently in a seemingly greater variability. A careful analysis of the data will not really show greater variability of judgment among the women. What it does show is that women have more and greater dislikes and are surer of them.

Hollingworth, however, used the method employed by Strong in chap. vi, and his results show women to be more variable than men by this very method. It is also true that to say that the

¹ E. K. Strong, Jr., *Relative Merits of Advertisements* (1911), pp. 78, 79.

women varied more *because* "they have more and greater dislikes, and are surer of them" is not to conclude that "a careful analysis of the figures will not show greater variability of judgment among the women." It is only to restate the fact that women *do* vary more in this case than men do, *in affective processes*.

Table XVI in Strong's monograph gives details from which he concluded that men are more variable. These figures show that the group of women does not differ as much from the first group of men in variability as the first group of men differs from the second group of men. For the group of women $Q=3.5$; for the first group of men $Q=4.0$; for the second group of men $Q=5.0$. Thus the group of women differs from the first group of men by .5, and from the second group of men by 1.5. Averaging these we get $d=1.00$. For the two groups of men $d=1.00$. On page 59 of his monograph Strong explains the great variability of the second group of men ($Q=5.0$) on the ground that the group is composed of uneducated persons who were possibly unable to differentiate complex appeals. Thus he explains a difference in variability between two groups of men on incidental grounds, but describes the same amount of difference in variability between a group of men and a group of women as a sex difference!

Gertrude Kuper¹ studied children of various ages and classes in their responses to a series of appeals. "The children numbered over 200, 10 boys and 10 girls for each year's age from 6.5 to 16.5. They were almost entirely attendants of the public schools of New York City, and came from quite varied sections of the city." This author draws the following conclusion:

A great sex difference was found in the variability measures as calculated for the various ages, appeals, social classes, and nationalities. In every case but two the girls exceeded the boys in their *P.E.*, and in these two exceptions the boys' *P.E.* was once greater than the girls' by 5 per cent, and another time exactly equal to the girls' *P.E.* The girls' average *P.E.* was 1.66; that for the boys was 1.36.

A monograph just published by Garry C. Meyers² offers an opportunity to note sex differences in variability, and is more

¹ Gertrude Kuper, "Group Differences in the Interests of Children," *Journal of Philosophy, Psychology, and Scientific Methods* (1912), p. 377.

² Garry C. Meyers, *Incidental Memory* (1913).

valuable from our point of view than any of the studies already cited, because he investigated a much greater number of subjects. His study of incidental memory of objects of common experience—bills, coins, and stamps—comprises 704 subjects—337 males and 367 females. Meyers classified these subjects into groups, and these groups range from third-grade pupils to college students, teachers, merchants, and bankers. The tables in which he gives the data for these groups separately have been studied and from them have been tabulated the number of groups of males showing greater variability than the corresponding group of females, and the number of groups of females showing greater variability than the corresponding group of males. The total number of groups is 182. Of this number 65 groups show greater variability for the males; 107 groups show greater variability for the females; 10 groups show exactly equal variability for both sexes. On the basis of these figures one might infer that females are much more variable than males. In his general conclusions about incidental memory for these objects Meyers himself says:

The amount of overestimation and underestimation of the sizes of the one dollar bill, stamp, and coins decreases as age and experience increases, and is as a rule greater for the females than for the males. Generally the males are better performers than the females, and less variable.

Meyers also studied incidental memory for words, using 1,663 subjects—773 males and 890 females. He states among his general conclusions:

The females are markedly superior to the males for average number of words remembered and for average efficiency; they have a high central tendency, vary more in the high schools and fourth grades; but in the fifth, sixth, seventh, and eighth grades they vary less than the males.

It must be noted here that the finding scarcely agrees with the exception previously quoted, i.e., that girls are more variable at the years nearest puberty, for on the average it seems likely that these two years would fall in the seventh and eighth grades, rather than in the fourth grade and the high school.

William Brown¹ in a study of *The Correlation of Mental Abilities* found that in groups of about equal homogeneity with respect

¹ W. Brown, "Correlation of Mental Abilities," *British Journal of Psychology* (1910), 296.

to age, training, etc., females are more variable in crossing out E, R; males are more variable in crossing out A, N, O, S; the sexes are equally variable in motor performance; males are more variable in the addition test, in speed, and females in accuracy; in the Müller-Lyer Illusion the male children are more variable, and the female adults are more variable.

Fox and Thorndike¹ studied arithmetical abilities of school children, using as subjects 28 boys and 49 girls. As to variability they conclude that in addition girls are only 93 per cent as variable as boys, and in multiplication only 96 per cent as variable.

Stone² also studied arithmetical abilities of school children in various school systems, using as subjects 250 girls and 250 boys. Six tests were given in four systems. Out of the 24 groups thus yielded, 9 show a greater variability for the boys, 14 show a greater variability for the girls, and 1 shows the same variability for both sexes. If we average the coefficients of variation for all groups, a procedure for which there seems to be little justification though not infrequently employed, the boys are found to be only 99.5 per cent as variable as the girls. Stone himself says:

This table shows that for the first two systems—the boys are somewhat more variable, and in systems 8 and 14 about the same amount less variable. This is interesting, and points to a need for further investigation, for the common opinion is that men are more variable than women; and supposedly boys more so than girls. But as seen by the averages for these four systems, so far as these 250 boys and 250 girls show the true tendency, there are no more exceptionally bright or exceptionally dull pupils among the boys than among the girls at this age.

Bonser³ in a study of arithmetical abilities of school children had a greater number of subjects than Stone and a much greater number than Fox and Thorndike. He tested 757 pupils—385 boys and 372 girls. He found that in arithmetical ability boys are only 66 per cent as variable as girls.

Bonser studied the reasoning ability of these 757 pupils with the result that in controlled association girls are once more variable

¹ Fox and Thorndike, "Sex Differences in Arithmetical Ability," *Columbia Contributions to Philosophy, Psychology, and Education*, XI.

² Stone, *Arithmetical Abilities* (1908), p. 36.

³ Bonser, *The Reasoning Ability of Children* (1910), p. 20.

and once less variable than boys; in selective judgment the girls are once more variable and once less variable than the boys; in arithmetical ability, as noted above, the girls are much more variable than the boys; in literary interpretation the boys are more variable; in spelling the boys are slightly more variable. Bonser's final conclusion regarding sex differences in variability in reasoning processes is as follows:

Taking the totals of all, the boys are slightly higher, the ratio being 1.047. The fluctuations are so numerous, and the differences so slight, that it seems unsound to make any general statement to the effect that the boys of these grades are more variable than the girls, in so far as these tests have shown.

Bonser's study affords a case¹ which illustrates very well the prime importance of considering the *whole table of frequencies* when we wish to infer social consequences. He distributed his subjects as to age, sex, and grade, and the medians and quartiles show much greater variability in age on the part of boys. Bonser states this fact as follows:

The variability in age is seen to be much greater among the boys than among the girls, as shown by a comparison of the Q's.

But fortunately for our purpose, Bonser gives the complete table of frequencies. From this we are able to see *in what* the greater male variability consists. We see that *the range* for the sexes is *equal*. At the *oldest extreme* we find 1.04 per cent of the boys and 1.88 per cent of the girls, while at the *youngest extreme* we find 0.51 per cent of the boys and 1.06 per cent of the girls. *The boys are more variable, but the highest achievements are more than twice as frequent, and the lowest achievements are nearly twice as frequent, on the part of girls.* The social significance would be the *exact opposite* of what greater male variability is ordinarily supposed to imply.

None of these studies was made for the chief purpose of studying sex differences in variability. The variations were calculated and stated more or less incidentally. There has been no attempt to select for reference here studies which found greater female variability. All studies known and accessible to the present writer,

¹ Bonser, *op. cit.*, p. 20.

where the variability of the sexes in mental traits has been computed, have been noted. In view of the facts that in many of the cases the conclusions are based on a small number of subjects, and that the evidence is conflicting, it seems necessary to conclude that the comparative variability of the sexes in mental traits has not been determined experimentally. If the evidence can be said to point in one direction rather than another, a greater female variability seems actually to be indicated in experiments so far made on the higher mental processes.

But even if it were determined that men *actually do* vary more in mental traits than women do, still nothing would be proved regarding their *inherent* variability. In order to establish the greater native variability of either sex it is necessary to show (1) that in the trait being distributed the opportunity and training of the sexes have been exactly equal, and (2) that in neither group has variability had more or less survival value than in the other group.

Under these conditions the only measurements of the sexes that may properly be compared with respect to variability are the measurements of infants at birth and for a short period thereafter. These are limited to anatomical traits, and objections are made to the validity of even these data. No measurements, especially mental measurements, of adults under the social customs which have obtained in the world of men and women fulfil either of our two necessary conditions. Men and women have devoted themselves to different activities because of the very different parts they play in the reproduction of the species. Women are under the biological necessity of bearing and rearing the children, and in the present almost as invariably as in the past, child-bearing has implied and compelled as a consequence the one occupation of housekeeping. Thus intellectual variability had no survival value for women, but rather the opposite. Women married, or were married by their parents, at an early age. They bore children—and *many* of them, since until the present century the very existence of a nation depended on the increase in its numbers of fighting men. All the influences of social pressure, religious precept, and even of the legal restriction of knowledge

have been brought to bear on women to the end that there might be enough increase in the population to offset the wastage of war and disease. Physiological facts made it natural, and consequent public expectation made it well nigh imperative, that women should contribute to the care of these numerous children by *housekeeping*. This was formerly almost absolutely the case, and even in this century the cases of women who have found a way to vary from the modal occupation and status, and yet procreate, are rare indeed. Individual prejudice hinders, poverty forbids, or society enacts legal measures against it, as in the case of a New York City teacher, which was recently given much publicity in the daily press. But men, except slave men, could always procreate and at the same time be as diverse in occupation, trade, and inclination as possible.

Thus (1) the opportunity and exercise of the two sexes in the traits which make for intellectual achievement have been very dissimilar in kind and amount, and (2) for one sex variability has had survival value; for the other sex it has had no survival value—this by virtue of the different parts played by the sexes in perpetuating the race. Darwin¹ says:

With respect to the causes of variability we are in all cases very ignorant, but we can see that in man, as in the lower animals, they stand in some relation to the conditions to which each species has been exposed during several generations. . . . We see the influence of diversified conditions in the more civilized nations; for the members, belonging to different grades of rank, and following different occupations, present a greater range of character than do the members of barbarous nations. . . .

This statement by Darwin involves, of course, a fallacy. For we do not know whether the civilized nations are more variable because they are civilized, or civilized because they are more variable. We can, however, paraphrase this statement and apply it to the situation of the two sexes. Men have been influenced by diversified conditions; they have followed the greatest possible range of occupations, and have at the same time procreated unhindered. Women have been limited to *one* set of activities, *because* of the part they play in the perpetuation of the species.

Men of science studying the ever-interesting subject of genius and leadership have pointed out women's inferiority to men in

¹ Charles Darwin, *Descent of Man* (1871), p. 44.

art, science, war, politics, and invention. They have diligently sought to explain the causes of this failure on the part of women. Ellis finds the causes in the greater primitiveness, less variability, and greater affectability of women. Lafitte finds the cause in the fact that women's minds are concrete and incapable of abstraction. Upton finds it in the fact that woman "is emotional by temperament and nature, and cannot project herself outwardly." Thorndike finds it chiefly in the greater variability of the male, and partly in the fact that women lack the fighting instinct. Countless men have found it in the "less ability" of women. None, so far as I know, has announced that he finds it in the conditioning influence of woman's biological function, the inescapable fact that she bears and rears the children. Frederic Harrison among general writers, in an essay on "The Future of Woman," recognizes the great influence that excessive maternity has had on woman's achievement:

We look to the good feeling of the future to relieve women from the agonizing wear and tear of families far too large to be reared by one mother—a burden which crushes down the best years of life for so many mothers, sisters, and daughters—a burden which, while it exists, makes all expectation of superior education or greater moral elevation in the masses of women mere idle talk.

Yet Harrison ends by forgetting this entirely, finding the final causes of woman's inferior achievement in "slighter nervous organization," "smaller cerebral mass," and in the fact that she is subject to the catamenial function and men are not.

J. McK. Cattell¹ in his study of the thousand most eminent persons of history says:

I have spoken thoughtout of eminent men as we lack in English words including both men and women, but as a matter of fact women do not have an important place on the list. They have in all 32 representatives in the thousand. . . . Belles lettres and fiction—the only department in which women have accomplished much—give ten names. . . . Women have not excelled in poetry or art. Yet these are the departments . . . in which the environment has been, perhaps, as favorable for women as for men. Women depart less from the normal than men—a fact that usually holds throughout the animal series. . . . The distribution of women is represented by a narrower, bell-shaped curve.

¹ J. McK. Cattell, "Statistical Study of Eminent Men," *Popular Science Monthly*, LXII.

It is interesting to notice that the "only department in which women have accomplished much" is one in which work could be carried on more or less successfully in conjunction with the modal occupation—providing there was wealth enough to hire servants for the actual drudgery. Cattell does not say explicitly what he means by the implied unfavorableness of the environment for women in lines other than art and poetry. He is not entirely certain that the environment has been as favorable for them as for men even in art and poetry, since he qualifies his statement by "perhaps." But it is clearly implied that this author recognizes an *environmental* condition unfavorable to women.

It seems indubitable that great numbers of women of intellectual gifts, confronted with the necessity of choosing a "career" or "domestic happiness," have chosen, either consciously or unconsciously the latter. And it must be remembered that even the possibility of a *choice* has existed only in recent times; that throughout almost the whole course of history women were predestined to their work of housekeeping. It is not and cannot be known how much nor what grade of potential leadership has thus been turned into energy-absorbing channels where eminence is impossible. Housekeeping and the rearing of children, though much commended to women as proper fields for the exploitation of their talents, are, unfortunately for their fame, not fields in which eminence can be attained. No one knows, for instance, who at present is the best housekeeper in America, nor who has borne and reared the largest and finest family of children. It is not known how much intellectual acumen is being brought to bear on these ends. Eminent housekeepers and eminent mothers *as such* do not exist. Yet to say that women of great intellectual gifts have not thus expended their energies is to affirm either (1) that there are no women of intellectual gifts, an affirmation now *passé* in the scientific world, (2) that intellect is unattractive to men, and that thus the most intelligent women are left unmarried, (3) that the most intelligent women will not marry, or (4) that the bearing and rearing of children, and the performance of household tasks at present coincident therewith constitute no handicap to the highest attainment in the fields where eminence is possible.

Such statements as these are very likely to be construed as an attack on maternity as such. It is certain, however, that no such attack is intended. The whole and the sole purpose of this paper is to criticize the hypothesis that inherently greater male variability is the cause of woman's failure to attain intellectual eminence. Such a criticism involves the unsentimental statement of biological facts, and of their social consequences. Men of science, seeking the cause of woman's failure, have not sufficiently recognized these facts and consequences, or else they have deemed it unpedagogical to announce them. We do not need, even, to look to the high ranges of achievement for light on our thesis. We need only to take the grade of intellectual attainment represented by the Ph.D. degree. It is proposed soon to make a comprehensive study of the percentage of women who have taken this degree after becoming mothers, as compared with the percentage of men who have taken it after becoming fathers. It is likely that any person of academic experience would forecast the result that few or no women have taken this degree after becoming mothers.

Cora Sutton Castle^{*} in her study of eminent women has attempted to determine why women have not played a greater part in the history of intellectual progress. She has treated eminent women with respect to their matrimonial relations, occupations, ages, nationalities, and epochs. But she has not yet determined *the number of children* borne by those women who attained eminence through *their intellectual labor*, as compared with the birth rate among women in general during the time when these women lived. Castle implies that woman's failure may be due to lack of educational opportunities, but we have farther to seek than that. For how did it come about that woman lacked educational opportunities? What was the genesis of this situation, since in the beginning there was no "educational opportunity" for either sex?

Thorndike has gone farther than almost any other man of science in declaring that woman's failure may to some extent be due to a difference in instincts *connected with reproduction*. He declares also that "We should first exhaust the known physical causes"

^{*} Cora Sutton Castle, *Statistical Study of Eminent Women* (1913).

before we proceed to any assumption of mental inferiority in explaining woman's lack of achievement. But have these "known physical causes" been exhausted if we end with the conclusion that "the probably true explanation is to be found in the greater variability within the male sex"? Surely we should consider *first* the established, obvious, inescapable, physical fact that women bear and rear the children, and that this has always meant and still means that *nearly 100 per cent of their energy is expended in the performance and supervision of domestic and allied tasks, a field where eminence is impossible*. Only when we had exhausted this fact as an explanation should we pass on to the question of comparative variability, or of differences in intellect or instinct. Men of science who discuss at all the matter of woman's failure should thus seek the cause of failure in the most obvious facts, and announce the conclusion consequent upon such search. Otherwise their discussion is futile scientifically.

Undoubtedly one of the most difficult and fundamental problems that today confront thinking women is how to secure for themselves the chance to vary from the mode of their sex, and at the same time to procreate, in a social order that has been built up on the assumption that there is and can be little or no variation in tastes, interests, and abilities within the female sex. It is a problem that has never confronted men. At times it seems well-nigh insoluble. But to affirm that it is insoluble is at the same time to affirm that there will always be a hard choice confronting women whose tastes vary from the mode; that there will be restlessness, unhappiness, and strife with the social order on the part of these individuals; and that society must tend to lose the work of its intellectual women or else lose their children.

Briefly our thesis may be summed up thus:

1. The greater variability of males in anatomical traits is not established, but is debated by authorities of perhaps equal competence.

2. But even if it were established, it would only suggest, not prove, that men are more variable in mental traits also. The empirical data at present available on this point are inadequate and contradictory, and if they point either way, actually indicate greater female variability.

3. But even if it were established that there *actually* is greater male variability in mental traits, it would only suggest, not prove, that there is greater *inherent* variability. For (a) the opportunity and exercise of the sexes have been dissimilar and unequal; (b) intellectual variability has had survival value for men, but for women it has had little or none—this by virtue of the different parts played by the sexes in the perpetuation of the species.

4. It must be remembered that variability in and of itself does not have social significance, unless it is known *in what* the variability consists—whether in greater range, greater frequency at the extremes, or in flattening at the top of the curve of distribution.

5. It is undesirable to seek for the cause of sex differences in eminence in ultimate and obscure affective and intellectual differences until we have exhausted as a cause the known, obvious, and inescapable fact that women bear and rear the children, and that this has had as an inevitable sequel the occupation of house-keeping, a field where eminence is not possible.

As a corollary it may be added that

6. It is desirable, for both the enrichment of society and the peace of individuals, that women may find a way to vary from their mode as men do, and yet procreate. Such a course is at present hindered by individual prejudice, poverty, and the enactment of legal measures. But public expectation will slowly change, as the conditions that generated that expectation have already changed, and in another century the solution to this problem will have been found.

BIBLIOGRAPHY CONSULTED

- Bonser, *Reasoning Ability of Children* (1910).
 Brown, W., "Correlation of Mental Abilities," *British Journal of Psychology* (1910).
 Castle, Cora S., *Statistical Study of Eminent Women* (1913).
 Cattell, J. McK., "Statistical Study of Eminent Men," *Popular Science Monthly*, 62.
 Darwin, Charles, *Descent of Man* (1871).
 Dubois, H., *Les gros enfants au point de vue obstetrical*. Thèse de Paris (1897).
 Ellis, H., *Man and Woman* (1909).
 Fox and Thorndike, "Sex Differences in Arithmetical Ability," *Columbia Contributions to Philosophy, Psychology, and Education*, II.
 Harrison, F., *Realities and Ideals* (1908).

- Hollingworth, H. L., "Judgments of Persuasiveness," *Psychological Review*, XXVIII (1911), 4.
- Hollingworth, L. S., "The Frequency of Amentia as Related to Sex," *Medical Record* (1913).
- Hrdlicka, A., *Anthropological Investigation of One Thousand Children* (1899).
- Kuper, G., "Group Differences in the Interests of Children," *Journal of Philosophy, Psychology, and Scientific Method* (1912).
- Meyers, G. C., *Incidental Memory* (1913).
- Pearson, K., *Chances of Death* (1897).
- Scripture, E. W., "Arithmetical Prodigies," *American Journal of Psychology* (1891).
- Stone, *Arithmetical Abilities* (1908).
- Strong, E. K., Jr., *Relative Merits of Advertisements* (1911).
- Thorndike, E. L., *Educational Psychology* (1910).
- Thorndike, E. L., "Sex in Education," *The Bookman*, XXIII, 213.
- Wells, F. L., "Sex Differences in the Tapping Test," *American Journal of Psychology* (1909).
- Wells, F. L., "Relation of Practice to Individual Differences," *American Journal of Psychology* (1912).